

Bioactivity exposure ratio (BER) analysis using high throughput transcriptomics (HTTr) chemical screening data

Joshua A. Harrill, USEPA National Center for Computational Toxicology (NCCT)



Application of non-animal approaches for decision-making in chemical safety assessment
NC3R, London, UK
December 10th, 2018

Disclaimer

The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or products represent endorsement for use.

Outline

- **Background**
 - Who is NCCT?
 - What does NCCT Do?
- **High-Throughput Transcriptomics (HTTr)**
 - Technology Overview
 - Experimental & Computational Workflows
 - Concentration-Response Modeling
 - Putative MIE/MOA Prediction with HTTR
- **Potential Applications for Regulatory Decision Making**
 - Bioactivity Exposure Ration (BER) Analysis

Who is NCCT?



Research Triangle Park Campus



National Center for Computational Toxicology

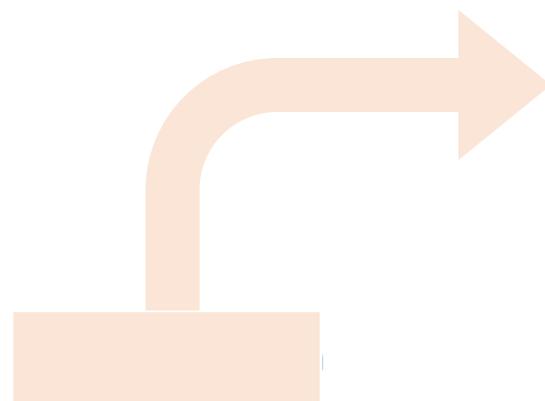


Mission Statement:

A research organization tasked with advancing the science of toxicity testing through the **development and/or application of novel experimental and computational approaches** for rapidly characterizing the biological activity, exposure potential and potential human health risks associated with chemicals.

What Does NCCT Do?

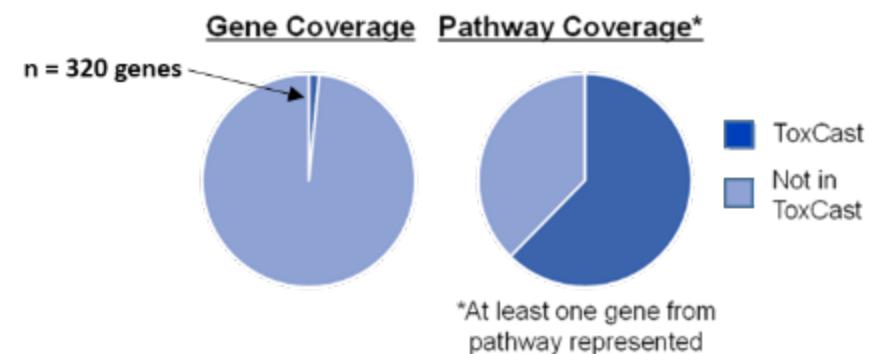
NCCT research programs focus on developing the **tools, approaches and data** needed to accelerate the pace of chemical risk assessment and foster incorporation of non-traditional toxicity testing data into regulatory decision-making processes.



- **ToxCast:** Use of high-throughput screening (HTS) assays to expose living cells or isolated proteins to chemicals and assess bioactivity and potential toxic effects.

	# of assays	# of chemicals	Types of chemicals
Phase 1 (2007 – 2009)	500	300	Mostly pesticides
Phase 2 (2009 – 2013)	700	2,000	Industrial, consumer product, food use, "green"

- Mostly targeted assays (*chemical X → protein Y*)
- Does not provide complete coverage of biological space

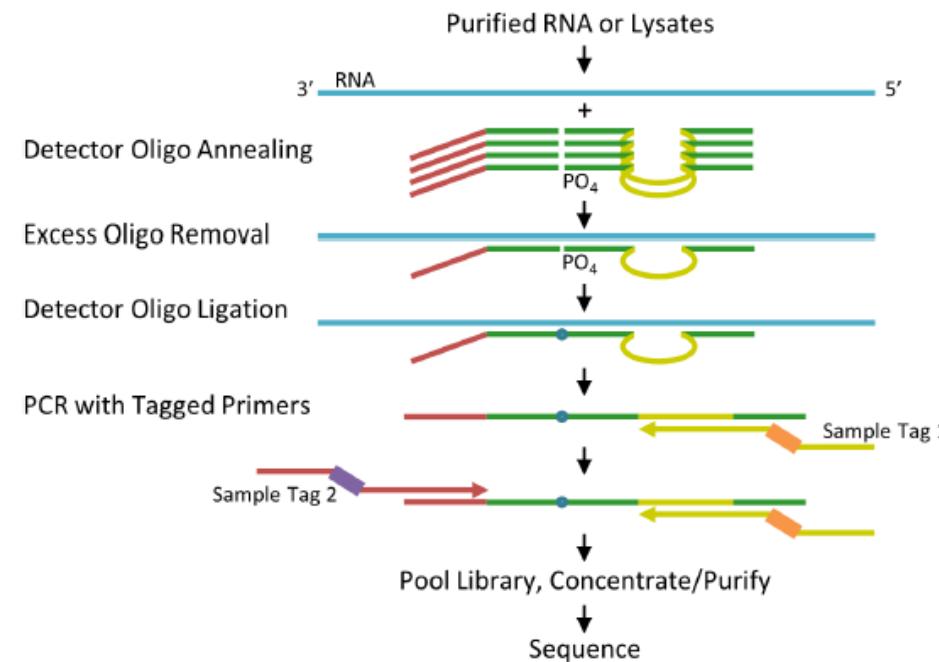


Templated Oligo with Sequencing Readout (TempO-Seq)

Technology

- The TempO-Seq human whole transcriptome assay measures the expression of greater than 20,000 transcripts.
- Requires only picogram amounts of total RNA per sample.
- Compatible with purified RNA samples or cell lysates.
- Transcripts in cell lysates generated in 384-well format are barcoded according to well position and combined in a single library for sequencing using industry standard instrumentation.
- Scalable, targeted assay:
 - 1) specifically measures transcripts of interest
 - 2) ~50-bp reads for all genes
 - 3) requires less flow cell capacity than RNA-Seq
- Per sample fastq files are generated and aligned to BioSpyder sequence manifest to generate integer count tables.

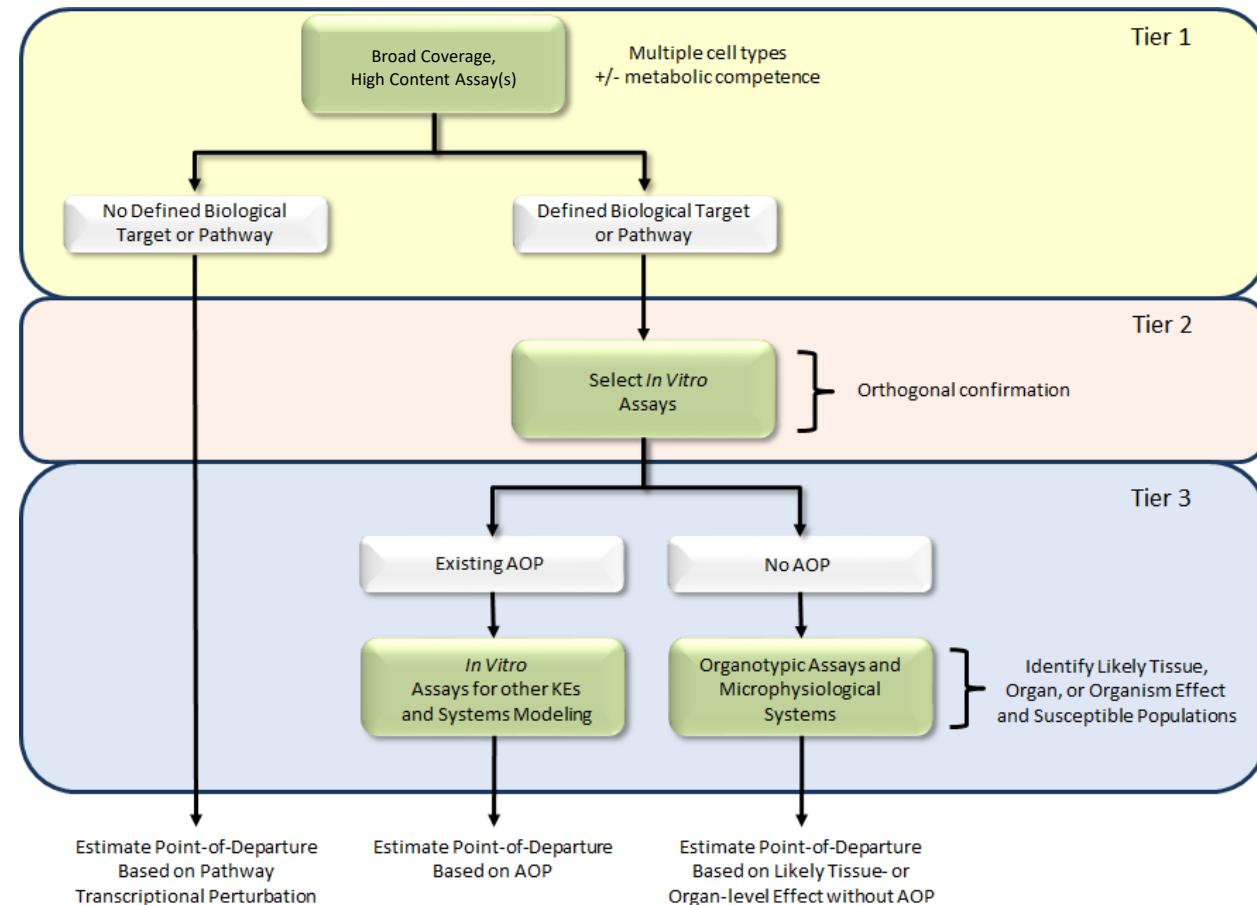
TempO-Seq Assay Illustration



Broad Coverage, High Content Screening Assays

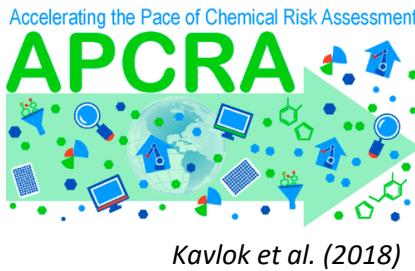
- Instead of targeted screening, NCCT proposes using screening strategies that cast the **broadest net possible** for capturing hazards associated with chemical exposure.
- Increasing efficiency and declining cost of generating whole transcriptome profiles has made **high-throughput transcriptomics (HTTr)** a practical option for broad coverage *in vitro* chemical screening.
- Bioactivity-based **potency estimates** can be used to identify *in vitro* **bioactivity thresholds**.
- Gene expression **profiles** can potentially be used for **mechanistic prediction** and evaluation of chemical similarity.

A Strategic Vision and Operational Roadmap for Computational Toxicology at US EPA



HTTr MCF-7 Screen: Experimental Design

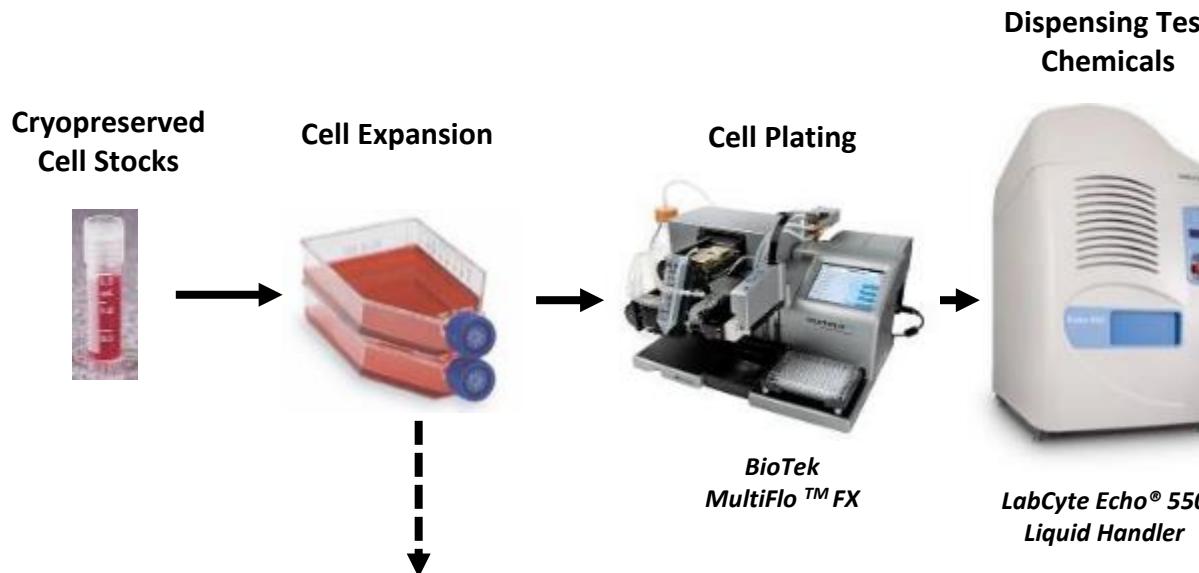
Parameter	Multiplier	Notes
Cell Type(s)	1	MCF-7
Culture Condition	1	DMEM + 10% HI-FBS ^a
Chemicals	2,112 (420)	ToxCast ph1, ph2, e1k / ph3 (APCRA)
Time Points:	1	6 hours
Assay Formats:	2	TempO-Seq HCl Cell Viability & Apoptosis
Concentrations:	8	3.5 log ₁₀ units; semi log ₁₀ spacing
Biological Replicates:	3	--



- International collaboration of regulatory scientists focused on developing case studies for evaluating the use of New Approach Methodologies (NAMs) in chemical risk assessment.
- ECHA Workshop (2017) case study focuses on **deriving quantitative estimates of risk based on NAM-derived potency information and computational exposure estimates**

EXPERIMENTAL AND COMPUTATIONAL WORKFLOWS FOR HTTR

Experimental Workflow

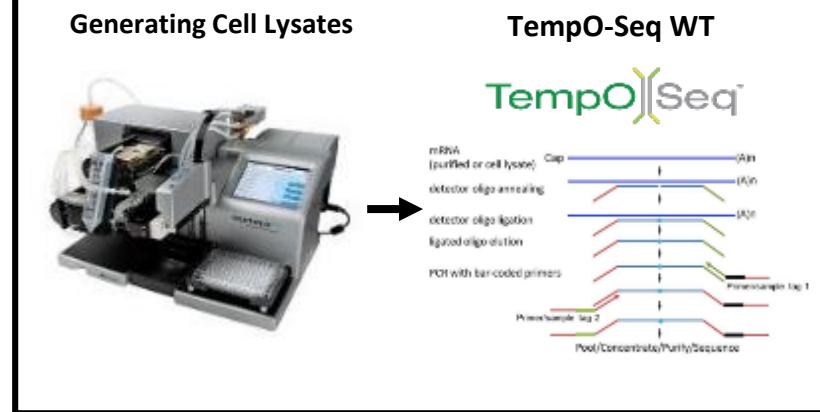


Standardized Expansion Protocol

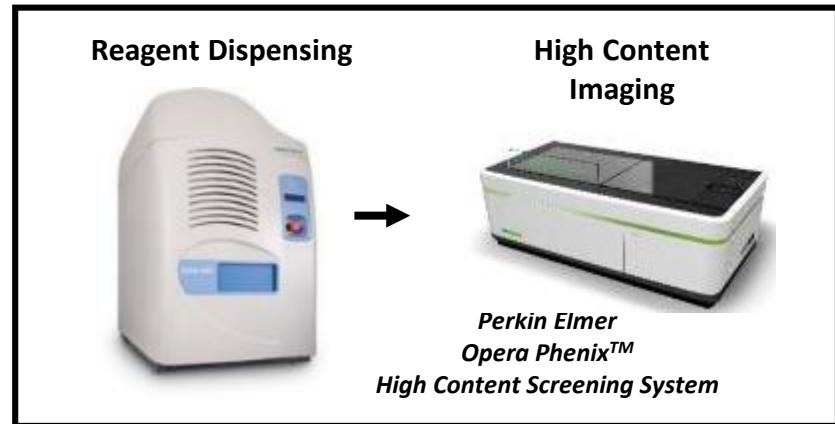
Day In Vitro (DIV):	0	2	5	7	9	11	13	
Action:	Seed	MC	P	MC	P	MC	P	MC = Media Change P = Passage
Vessel:	T25		T75		T225		Test Plate(s)	Perform Experiment

Dispensing Test Chemicals

Track 1: Targeted RNA-Seq

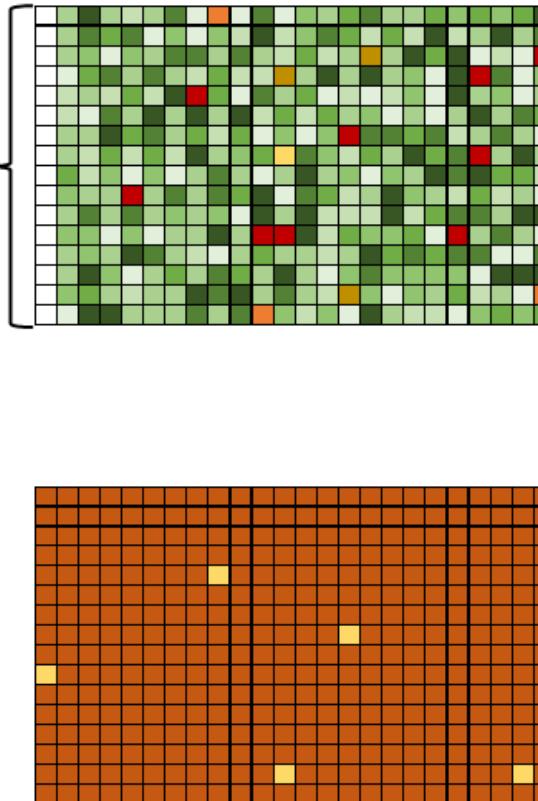
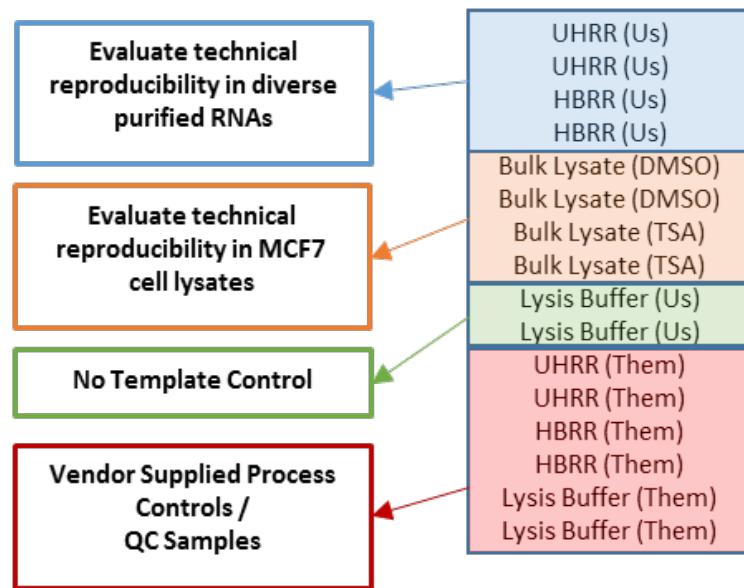


Track 2: Apoptosis / Cell Viability

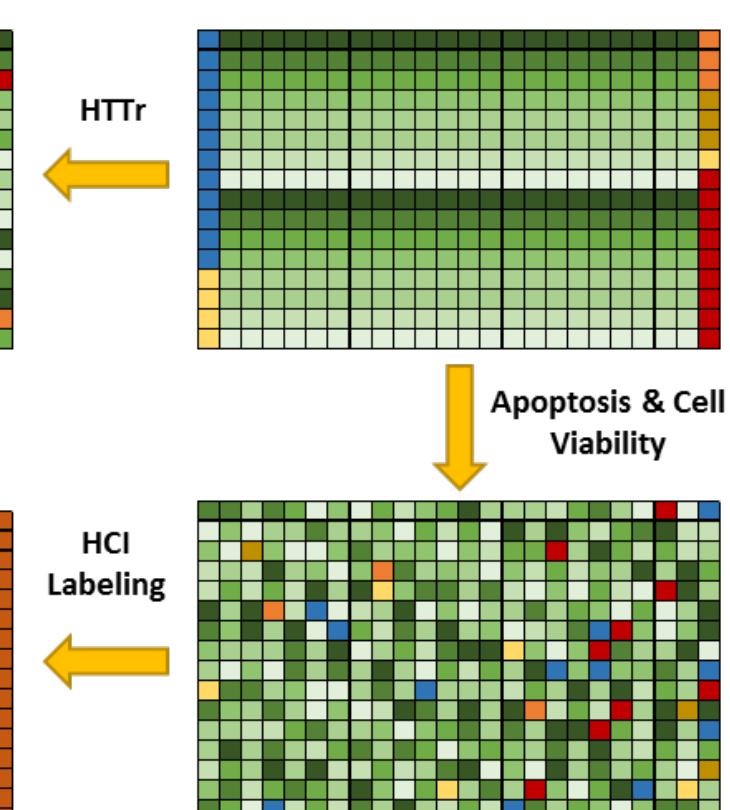


Treatment Randomization & Quality Control Samples

Treatment Randomization: *Each test plate uniquely randomized with respect to treatment.*
QC Samples: *Quality Control samples included on each plate*

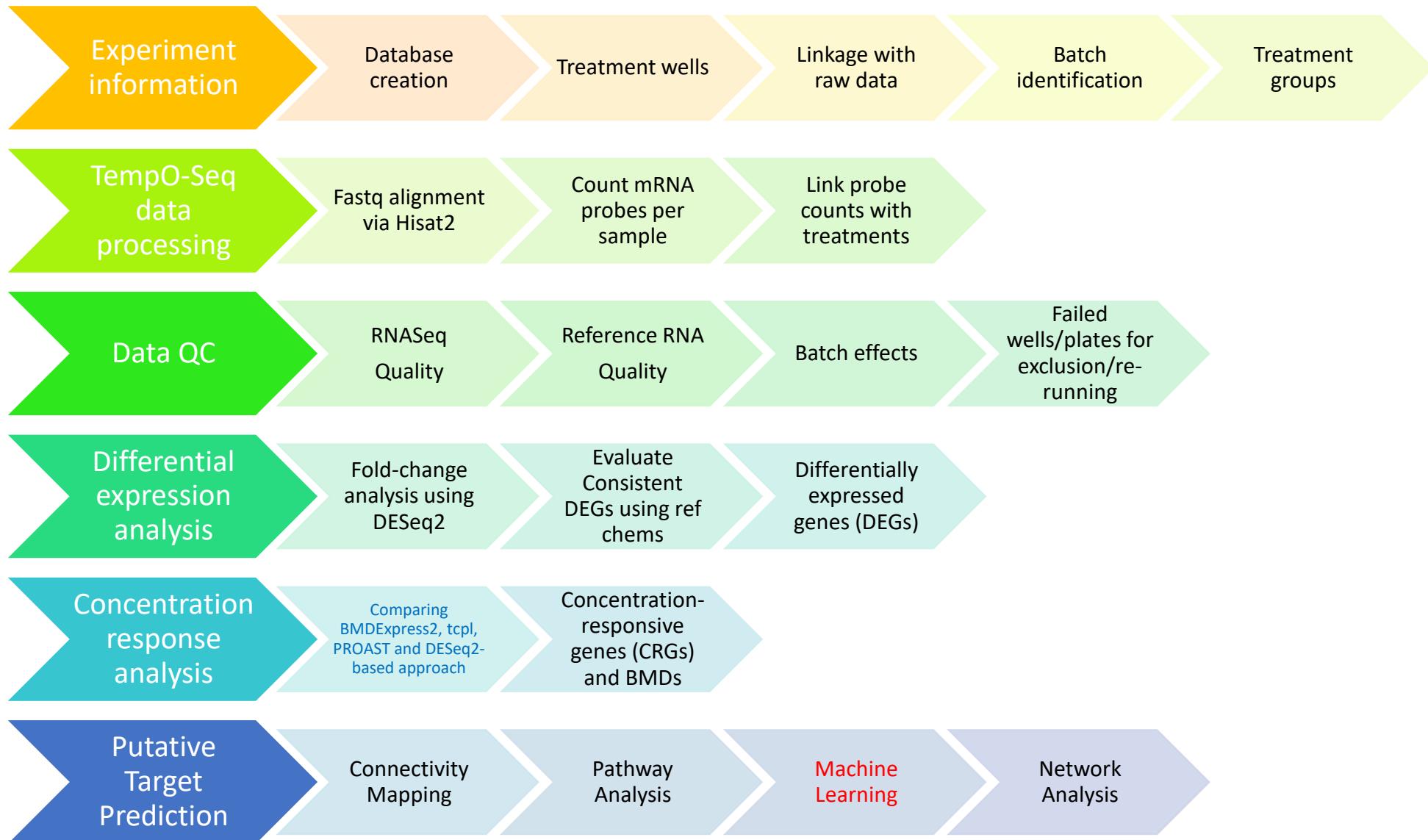


UHRR = Universal Human Reference RNA
 HBRR = Human Brain Reference RNA



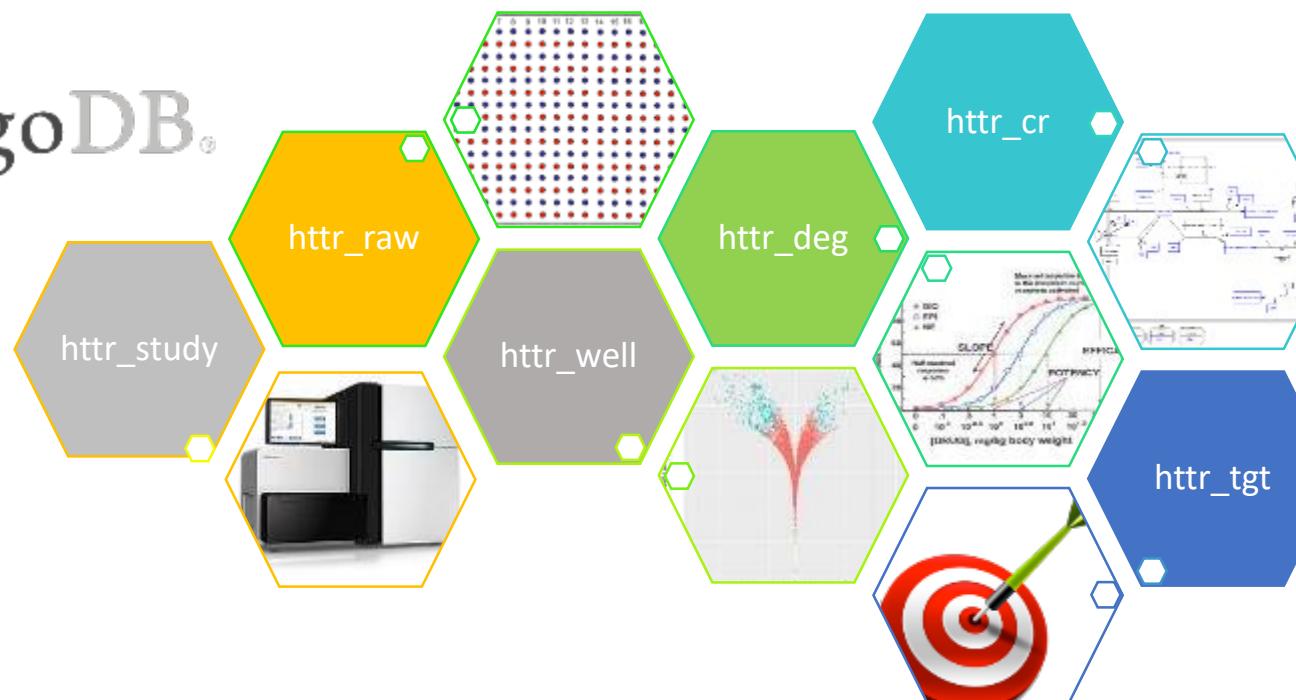
Test Chemicals
Untreated
DMSO (vehicle control)
CMPA Reference
HCI No Label Controls
HCI Pos. & Neg. Controls

HTTr Analysis Pipeline (Nov 2018)

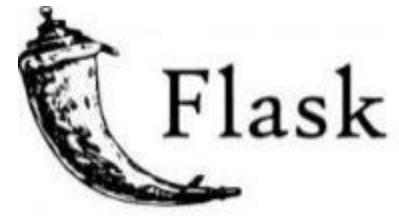


HTTr Computational Framework and Infrastructure

Python & R analysis pipeline



REST API



<http://httr-dev.epa.gov/api/httr/v1/>

searchChem
getChemPlates
getPlateInfo
getPlateGroups
getChemProbeCounts
getChemDEG

getChemCRG
getChemTargets

<http://bitbucket.zn.epa.gov/projects/HTTR>

Concentration-Response Modeling



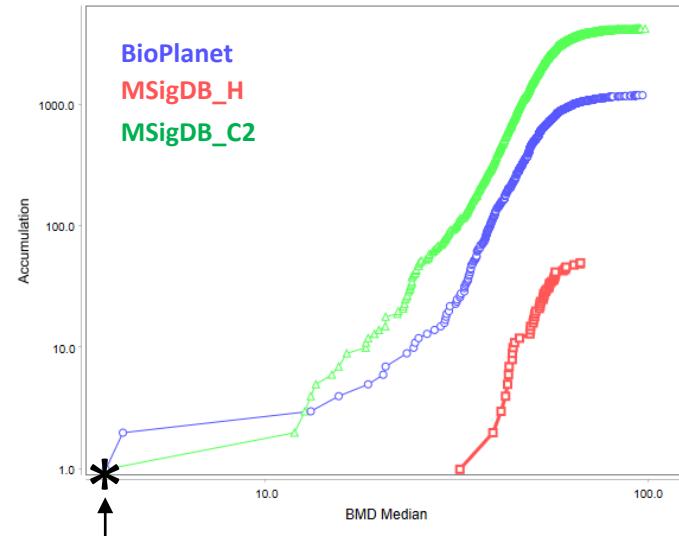
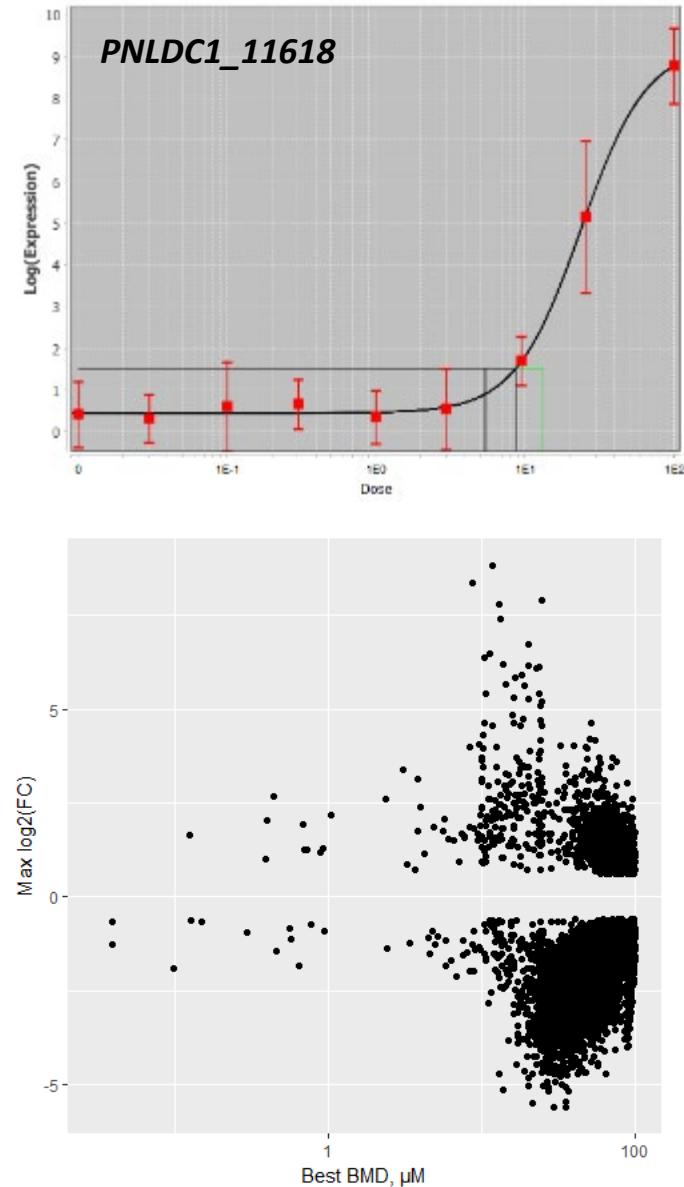
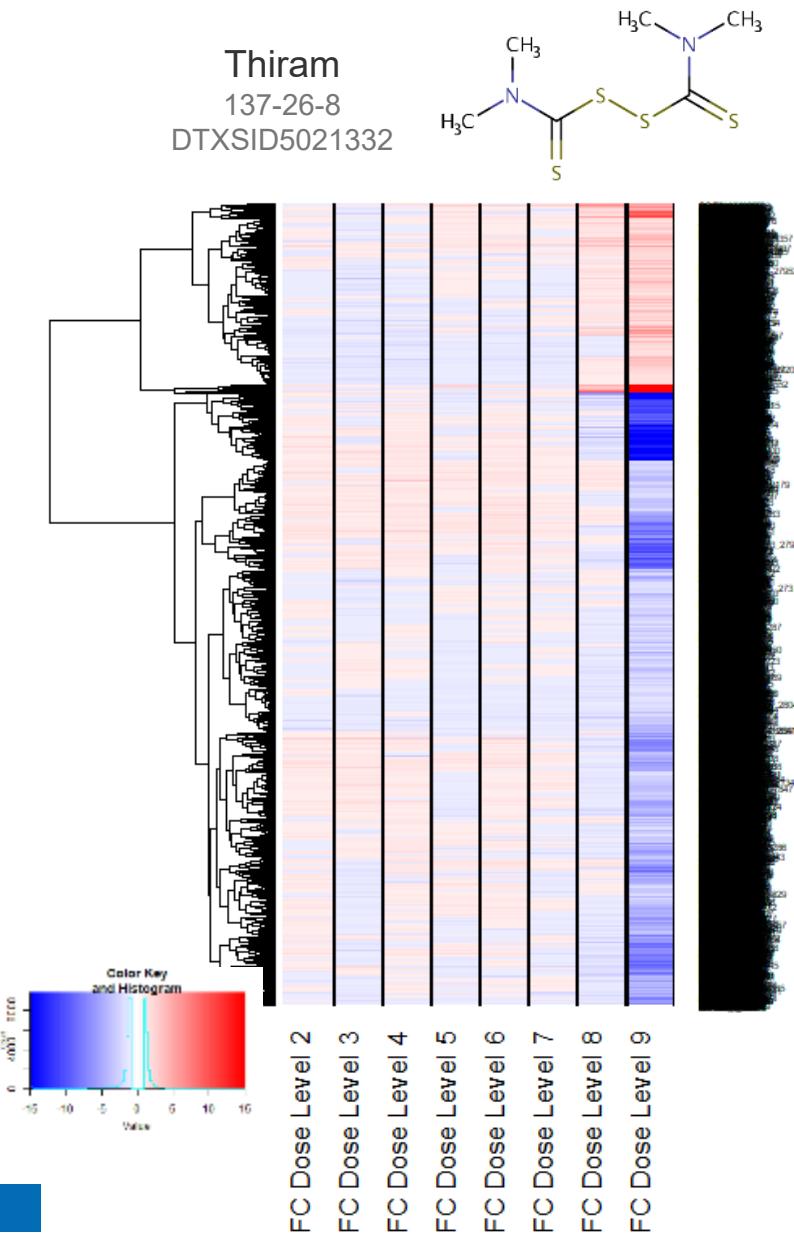
Parameter	Criteria ^a
Pre-filter:	William's Trend Test ($p_{\text{raw}} < 0.05$ & $ FC \geq 1.5$)
Models	Hill, Exponential 3, <i>poly2, power, linear</i>
BMR Factor:	1.349 (10 %)
Best Model Selection:	Lowest AIC Best Fit p-value > 0.1
Hill Model Flagging ^b :	'k' < 1/3 Lowest Positive Dose Retain Flagged Models
Pathway Analysis:	Genes with $\text{BMD} \leq \text{Highest Dose} \geq 3$ $\geq 5\%$ Gene Set Coverage
Gene Set Collections ^c :	Molecular Signatures Database BioPlanet

^c Gene Set Collections:

- **MSigDB_C2:** A collection of curated gene sets from publicly-accessible online resources including KEGG, BioCarta and Reactome (n = 4,725)
- **MSigDB_H:** Coherently expressed signatures derived by aggregating many MSigDB gene sets to represent well-defined biological states or processes (n = 50).
- **BioPlanet:** Curated set of non-redundant biological pathways developed by NIH-NCATs (n = 1,700).

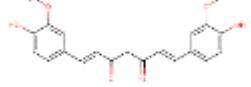
^a Exploratory analysis – modeling criteria not finalized

Concentration Response Modeling Example

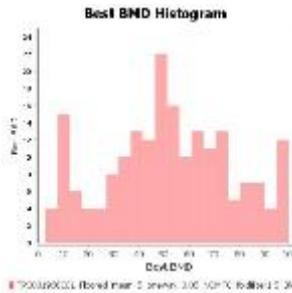


**Biological Pathway Altering
Concentration (BPAC)**

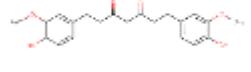
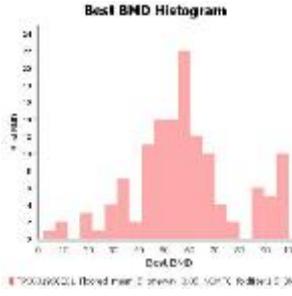
Gene Set Analysis Using BMD Modeling Results



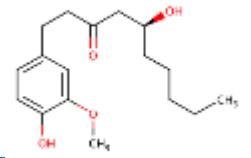
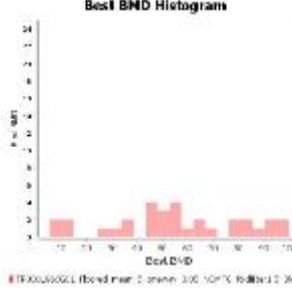
Curcumin | DTXSID0831077



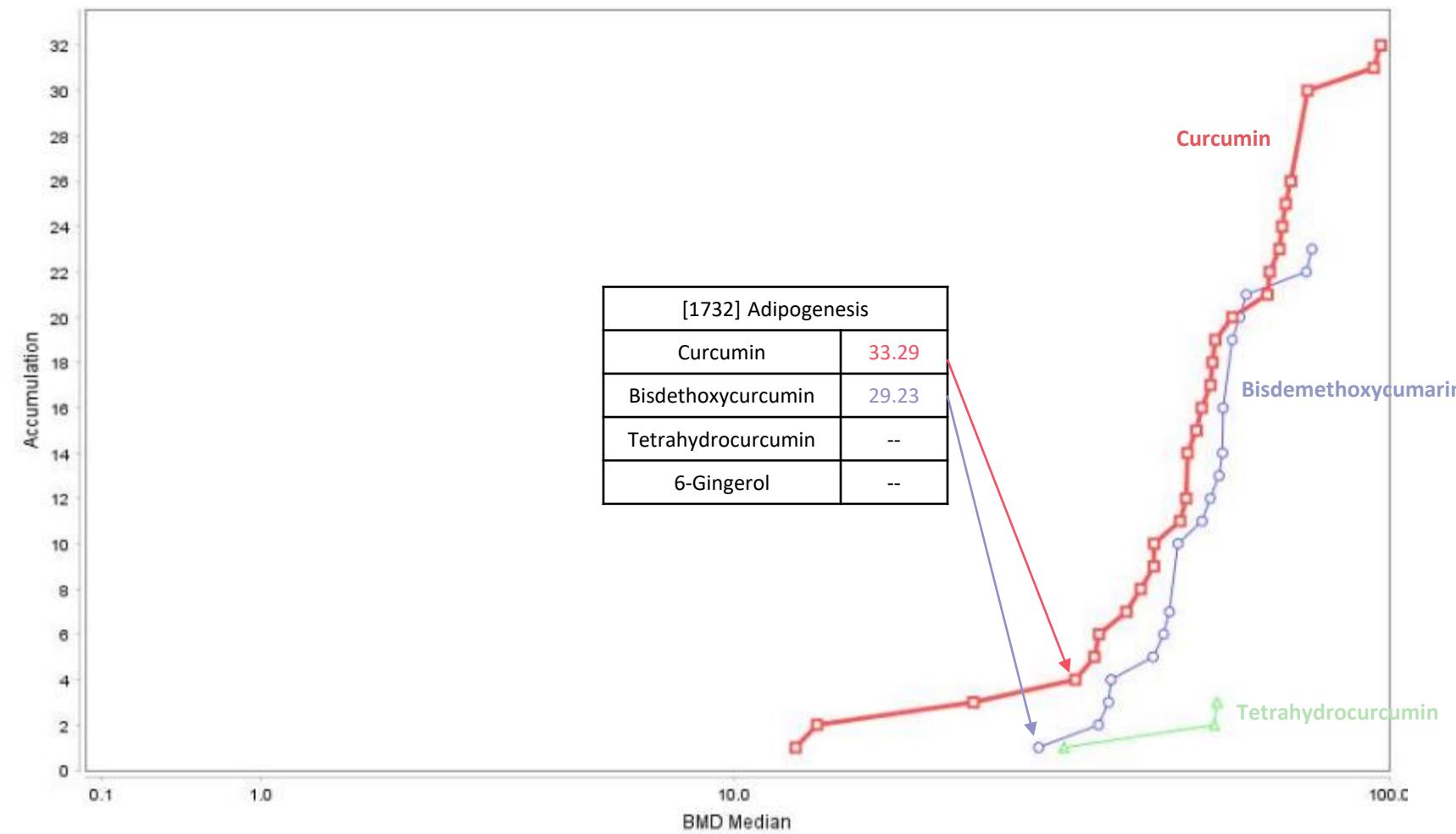
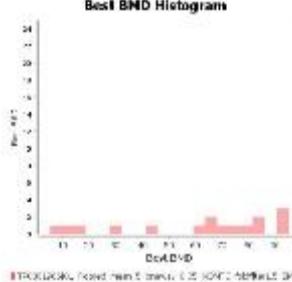
Bisdemethoxycurcumin | DTXSID00872663



Tetrahydrocurcumin | DTXSID30865801



(6)-Gingerol | DTXSID3041035



[Toxicol Appl Pharmacol.](#) 2017 Aug 15;329:158-164. doi: 10.1016/j.taap.2017.05.036.

Curcumin inhibits adipogenesis induced by benzyl butyl phthalate in 3T3-L1 cells.

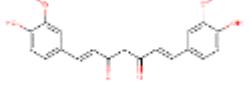
Sakuma S¹, Sumida M², Endoh Y², Kurita A², Yamaguchi A², Watanabe T², Kohda T², Tsukiyama Y², Fujimoto Y³.

[J Agric Food Chem.](#) 2016 Feb 3;64(4):821-30. doi: 10.1021/acs.jafc.5b05577. Epub 2016 Jan 25.

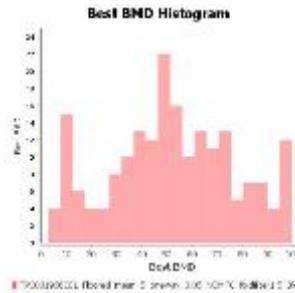
Bisdemethoxycurcumin Inhibits Adipogenesis in 3T3-L1 Preadipocytes and Suppresses Obesity in High-Fat Diet-Fed C57BL/6 Mice.

Lai CS^{1,2}, Chen YY¹, Lee PS¹, Kalyanam N³, Ho CT⁴, Liou WS⁵, Yu RC¹, Pan MH^{1,6,7,8}.

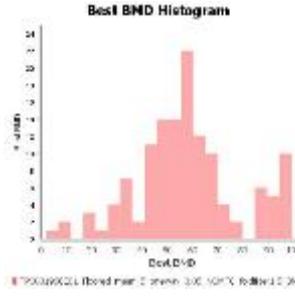
Gene Set Analysis Using BMD Modeling Results



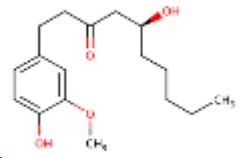
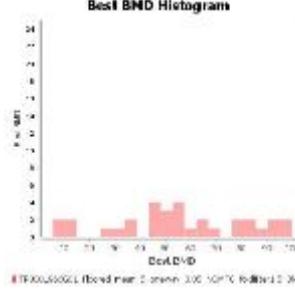
Curcumin | DTXSID8031077



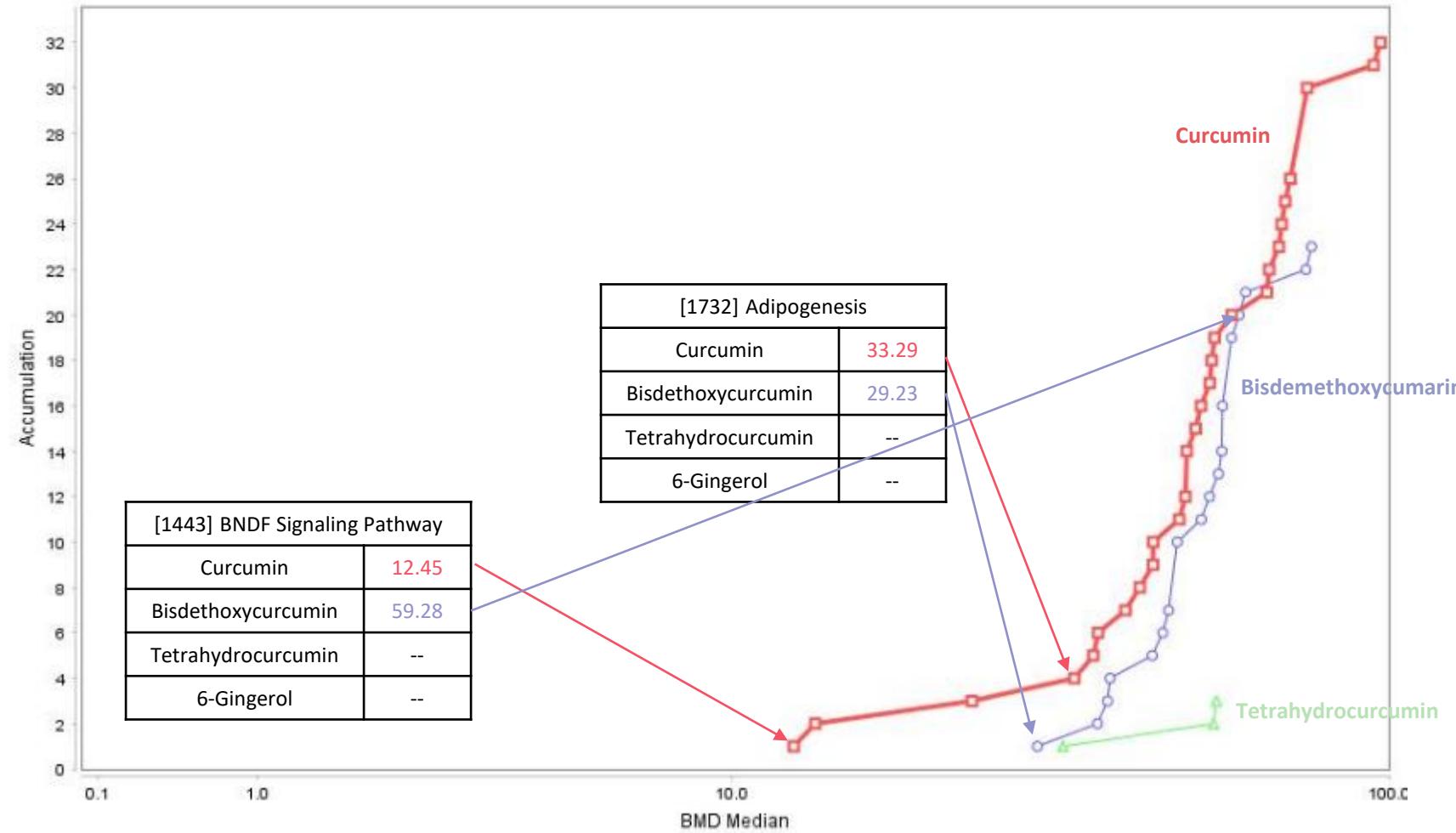
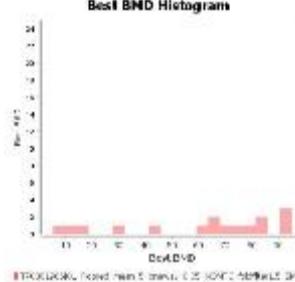
Bisdemethoxycurcumin | DTXSID00872663



Tetrahydrocurcumin | DTXSID30865801



(6)-Gingerol | DTXSID3041035



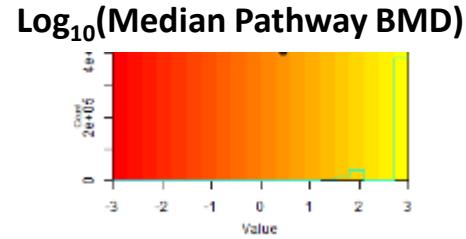
[Neuropeptides](#). 2016 Apr;56:25-31. doi: 10.1016/j.npep.2015.11.003. Epub 2015 Nov 11.

Effect of curcumin on serum brain-derived neurotrophic factor levels in women with premenstrual syndrome: A randomized, double-blind, placebo-controlled trial. [Fanaei H¹](#), [Khayat S²](#), [Kasaeian A³](#), [Javadimehr M⁴](#).

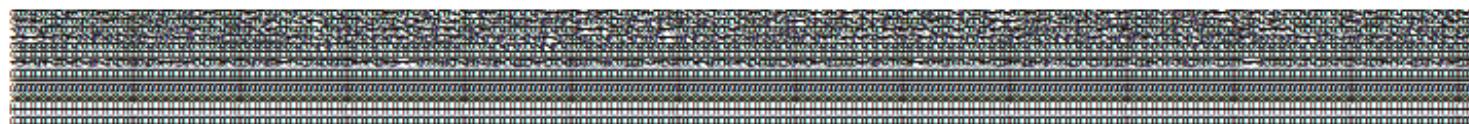
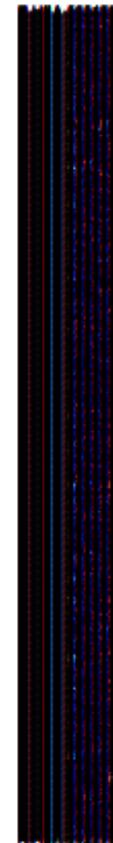
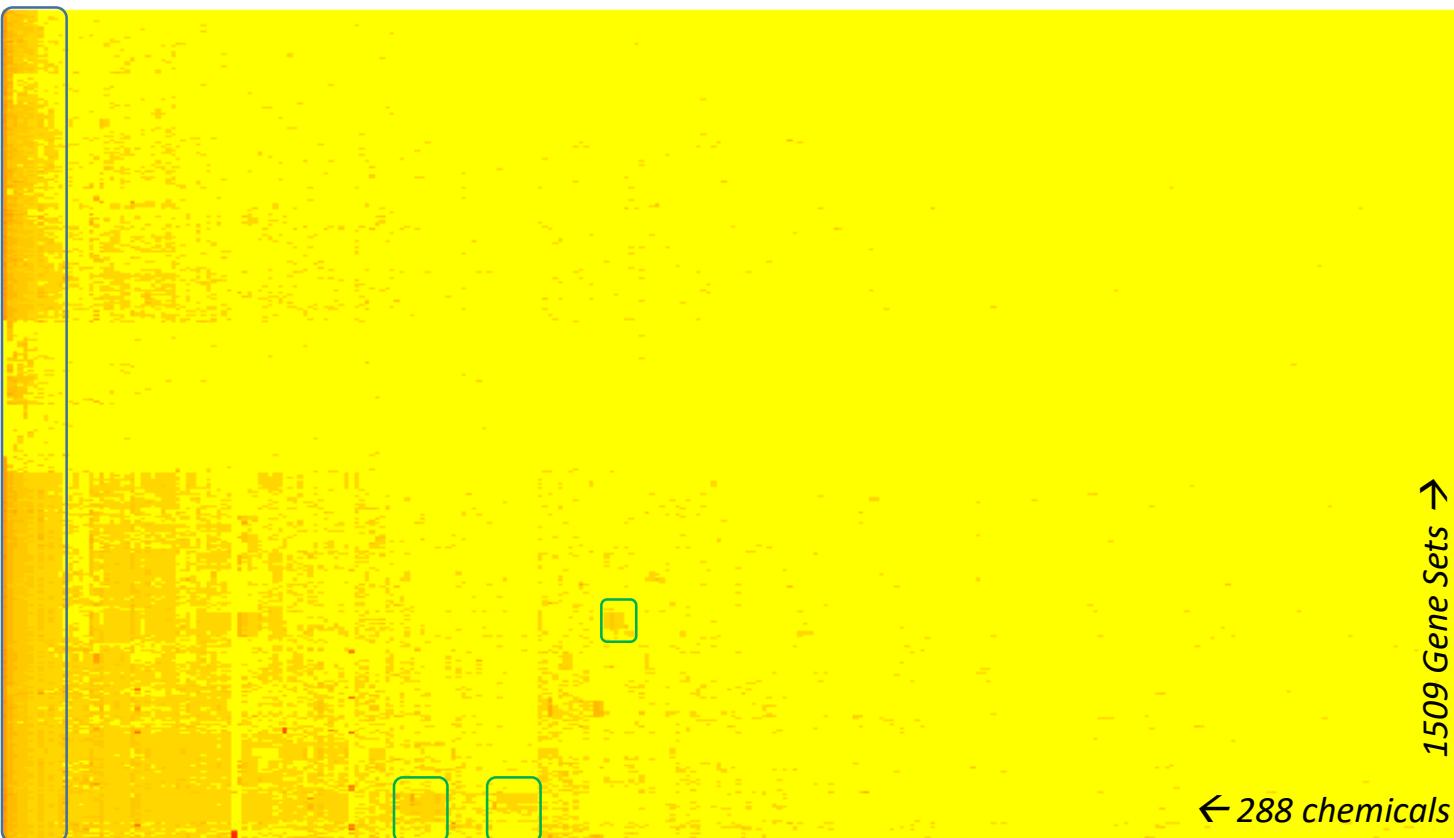
[Biomed Pharmacother](#). 2017 Mar;87:721-740. doi: 10.1016/j.biopha.2016.12.020. Epub 2017 Jan 14.

Curcumin confers neuroprotection against alcohol-induced hippocampal neurodegeneration via CREB-BDNF pathway in rats. [Motaghinejad M¹](#), [Motevalian M²](#), [Fatima S³](#), [Hashemi H¹](#), [Gholami M⁴](#).

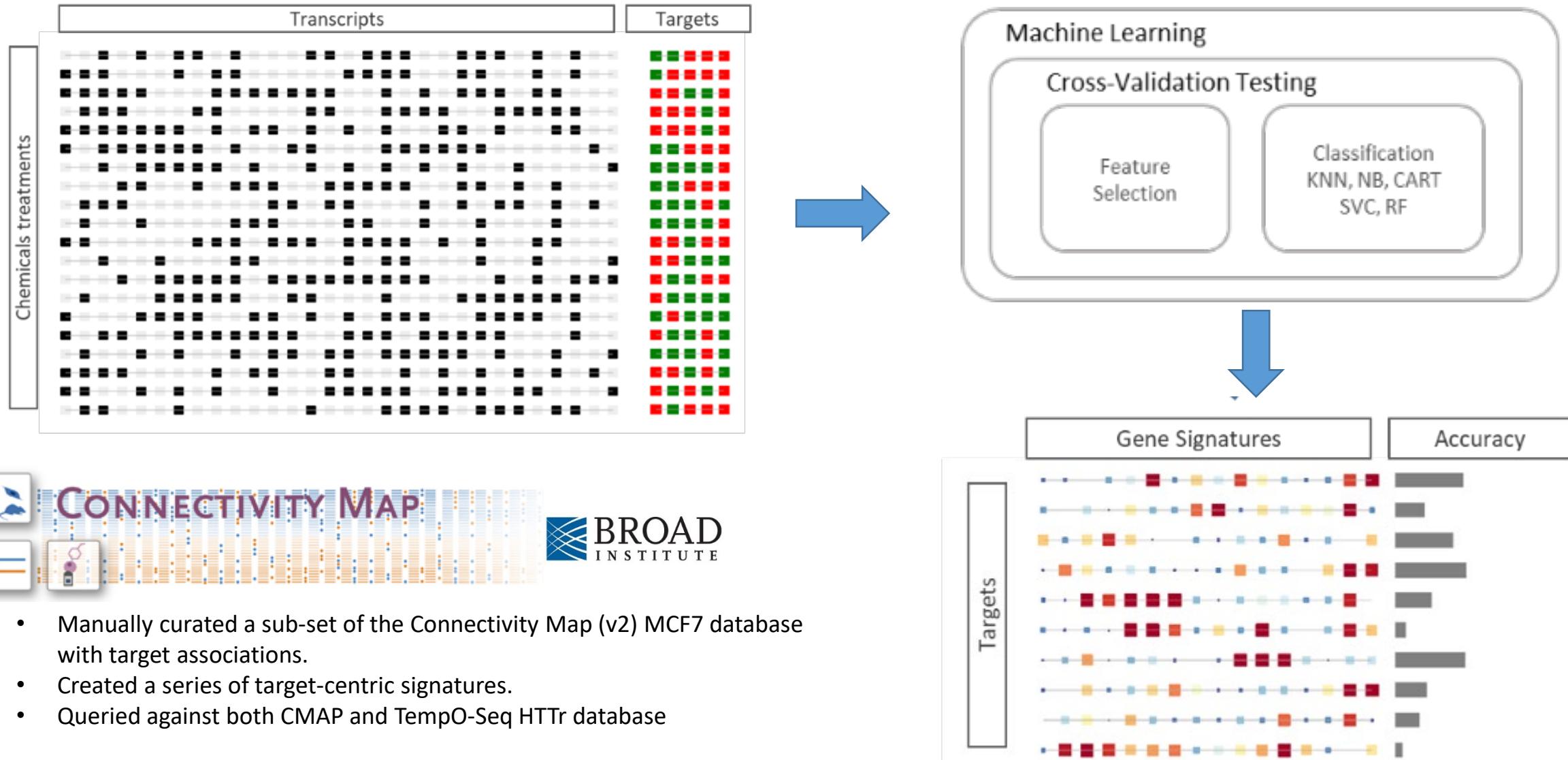
Qualitative Similarities in Gene Expression Profiles



Observe both **promiscuous** chemicals
and **profile similarities** across chemicals



Signatures/Classifiers For Putative Target Prediction



HDAC Inhibitors Derived from CMAP

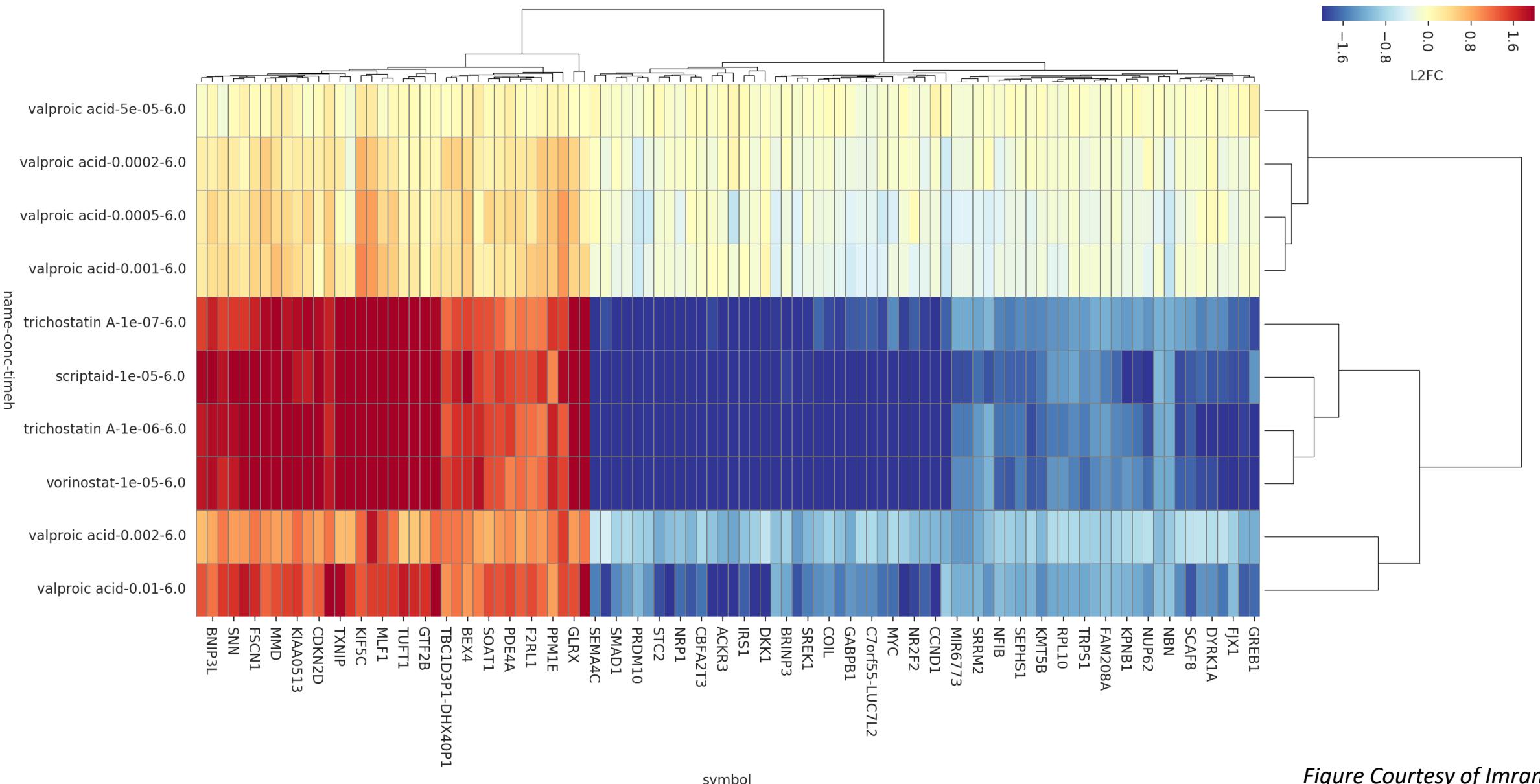


Figure Courtesy of Imran Shah

ER Model (any Mode) Derived from CMAP

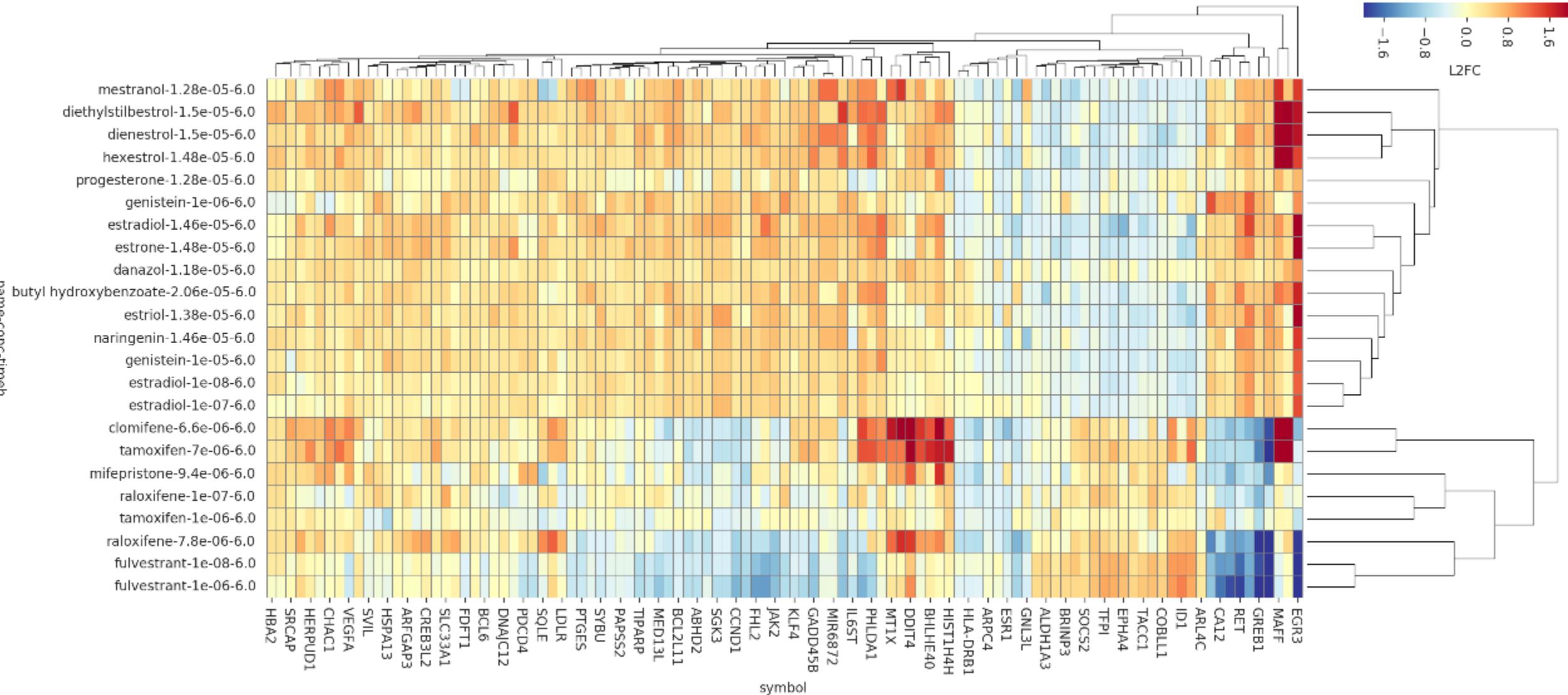


Figure Courtesy of Imran Shah

AR Model (any Mode) Derived from CMAP

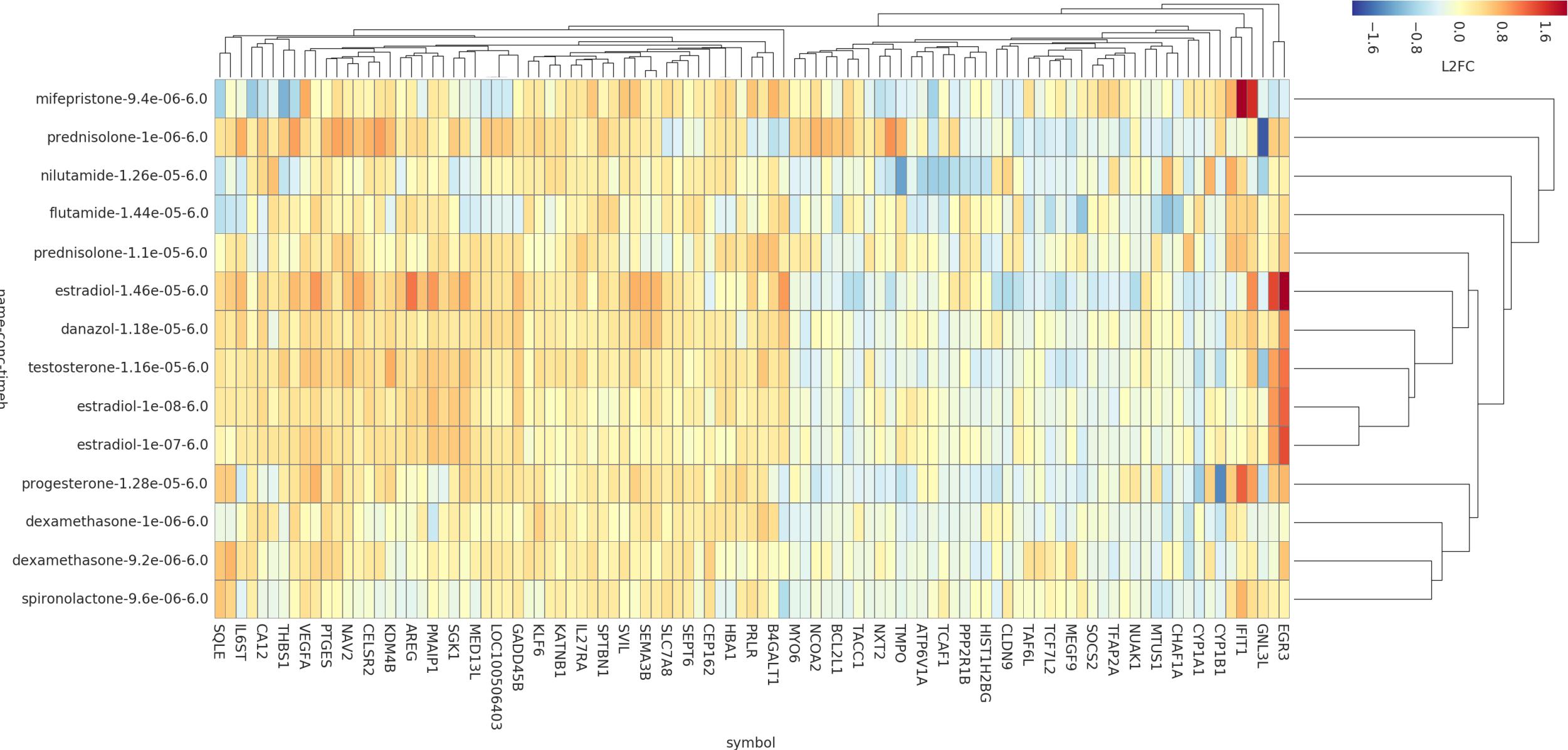


Figure Courtesy of Imran Shah

NR3C1 (Glucocorticoid Receptor) Derived from CMAP

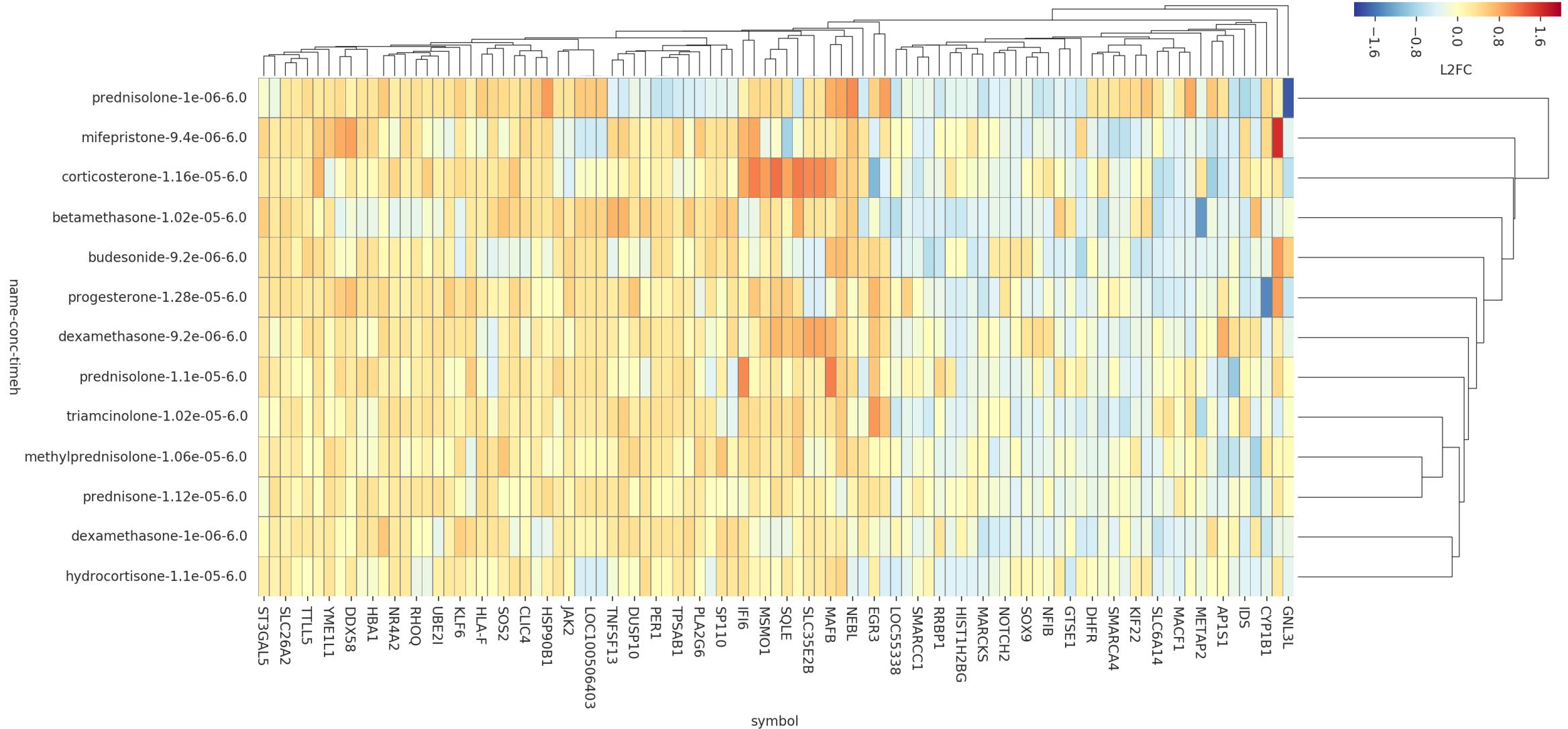


Figure Courtesy of Imran Shah

Searching CMAP v2 Using ER Signature

dsstox_sid	name	target	cell	conc	jaccard	olap	timeh
Other	equilin	Other	MCF7	0.000015	0.176	110	6
Other	equilin	Other	MCF7	0.000015	0.166929	106	6
Other	lynestrenol	Other	MCF7	0.000014	0.163303	89	6
Other	prasterone	Other	MCF7	1.22E-05	0.162055	82	6
DTXSID3020465	diethylstilbestrol	ESRRG ES	MCF7	0.000015	0.153846	82	6
DTXSID3020465	diethylstilbestrol	ESRRG ES	MCF7	0.000015	0.144531	74	6
Other	trifluoperazine	Other	MCF7	0.00001	0.140777	87	6
Other	epitiostanol	Other	MCF7	0.000013	0.137876	87	6
Other	equilin	Other	MCF7	0.000015	0.1376	86	6
Other	etynodiol	Other	MCF7	1.04E-05	0.135314	82	6
DTXSID0020814	mestranol	ESR1	MCF7	1.28E-05	0.134052	87	6
DTXSID8022371	testosterone	AR ESR2	MCF7	1.16E-05	0.131068	81	6
DTXSID6023656	thioridazine	DRD2 KCNQ5	MCF7	0.00001	0.127303	76	6
DTXSID6023656	thioridazine	DRD2 KCNQ5	MCF7	0.00001	0.126761	72	6
DTXSID4022369	fulvestrant	ESR2 ESR1	MCF7	1E-08	0.12623	77	6
Other	suloctidil	Other	MCF7	1.18E-05	0.124756	64	6
Other	prochlorperazine	Other	MCF7	0.00001	0.124214	79	6
DTXSID5022308	genistein	ESR2 ESR1	MCF7	0.00001	0.123664	81	6
Other	suloctidil	Other	MCF7	1.18E-05	0.120553	61	6
Other	mometasone	Other	MCF7	7.6E-06	0.119869	73	6
DTXSID2022880	danazol	AR ESR1	MCF7	1.18E-05	0.119449	78	6
Other	trifluoperazine	Other	MCF7	0.00001	0.11936	82	6
Other	fluphenazine	Other	MCF7	0.00001	0.119163	74	6
DTXSID9020110	astemizole	KCNH2 H1	MCF7	8.8E-06	0.119005	67	6
Other	butyl hydroxybenzoate	Other	MCF7	2.06E-05	0.118902	78	6
Other	ciclosporin	Other	MCF7	3.4E-06	0.118699	73	6
Other	ivermectin	Other	MCF7	4.6E-06	0.118098	77	6

“Horse Estrogen”

Synthetic progesterone

Synthetic progesterone

Pro-androgen / estrogen

Dopamine antagonist /
Antipsychotic
Gynecomastia in males

*Most of the top
hits are ER-related*

Performance of Signatures for Putative Target Prediction in HTTr Data

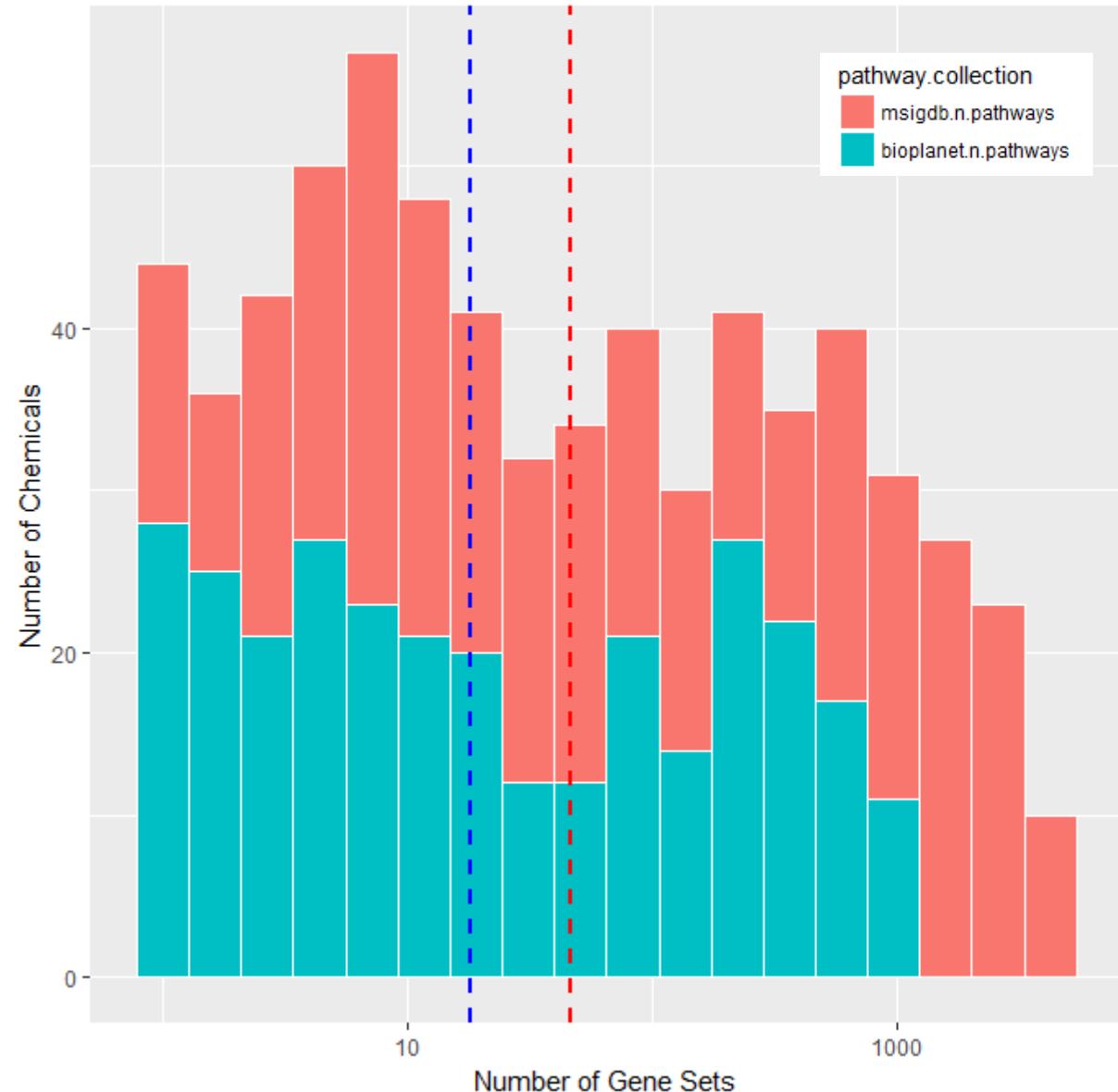
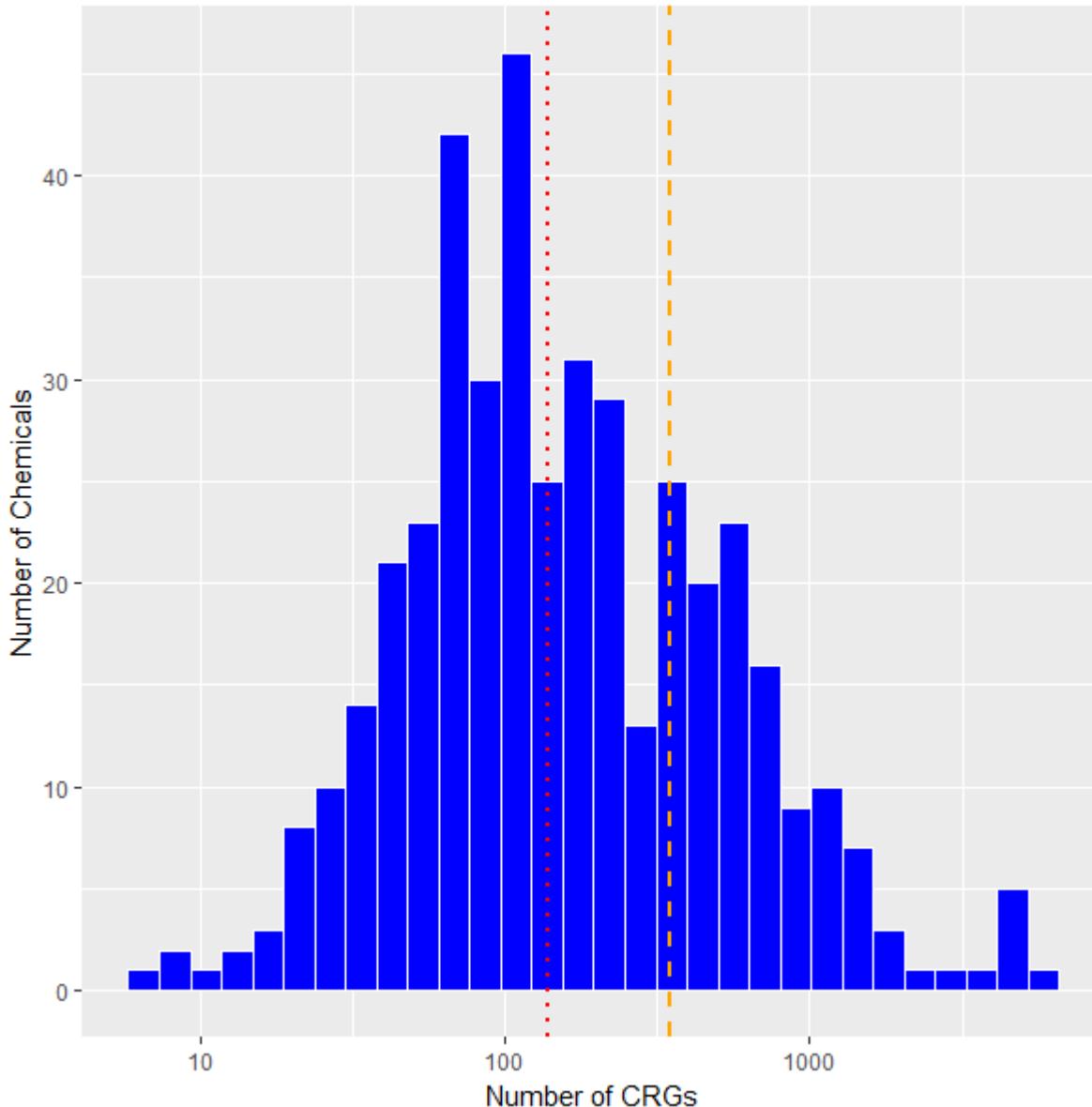
Slide by Imran Shah

Putative Target	CMap v2 / Affymetrix	BioSpyder HTTr-Phase I			
	Signature size	PPV	Positives	Positive Chemicals found (Curated)	Top 5 Prediction (Uncurated)
CYP2C9	131	1	1	Fluconazole	Emodin, Phenazopyridine hydrochloride, Lactofen, Hexachlorophene, 2-Amino-5-azotoluene
ESR1	257	1	11	o,p'-DDT, Genistein, 4-Nonylphenol, 4-Hydroxytamoxifen, Diethylstilbestrol, Raloxifene hydrochloride, Bisphenol A, 17beta-Estradiol, 5alpha-Dihydrotestosterone, Mifepristone, 4-(1,1,3,3-Tetramethylbutyl)phenol	dl-Norgestrel, SSR504734, Haloperidol, Cyclosporin A, Astemizole
HDAC1	124	1	2	Trichostatin A, Valproic acid	2-(Thiocyanomethylthio)benzothiazole, Azinphos-methyl, Sodium (2-pyridylthio)-N-oxide, 3,3'-Dichlorobenzidine dihydrochloride
DHFR	215	1	2	Pyrimethamine, Methotrexate	Adriamycin hydrochloride, PharmaGSID_48505, Etoposide, Resveratrol, Nisoldipine
NR1I2	139	1	2	17beta-Estradiol, Bisphenol A	dl-Norgestrel, Endosulfan, Isodrin, Genistein, 17alpha-Estradiol
PGR	115	1	1	Mifepristone	Flurandrenolide, Fluorometholone, Dexamethasone, Melengestrol acetate, Betamethasone
HMGCR	236	1	1	Lovastatin	Resveratrol, dl-Norgestrel, o,p'-DDT, Tamoxifen, Chlorhexidine
ABCC2	357	1	1	Methotrexate	4-Nitrosodiphenylamine, Resveratrol, Adriamycin hydrochloride, Nisoldipine, 8-Hydroxyquinoline sulfate
TYMS	329	1	1	Methotrexate	Etoposide, Resveratrol, 4-Nitrosodiphenylamine, Cytarabine hydrochloride, PharmaGSID_48505
ESR2	281	0.86	7	Genistein, Diethylstilbestrol, 4-Nonylphenol, Bisphenol A, 4-Hydroxytamoxifen, 17beta-Estradiol	dl-Norgestrel, 17alpha-Estradiol, Haloperidol, Cyclosporin A, Isodrin
AR	261	0.78	9	o,p'-DDT, 17beta-Estradiol, 5alpha-Dihydrotestosterone, Flutamide, Bisphenol A, Mifepristone, 17-Methyltestosterone	dl-Norgestrel, Melengestrol acetate, Dehydroepiandrosterone, 8-Hydroxyquinoline, Genistein
NR3C2	352	0.5	2	Mifepristone	Fluocinolone acetonide, Bexarotene, 1-Naphthol, Dexamethasone, dl-Norgestrel
ABCB1	117	0.5	2	Reserpine	Fabesetron hydrochloride, Abamectin, SAR115740, SSR69071, Chlorbenzilate
NR3C1	148	0.5	4	Triamcinolone, Mifepristone	Medroxyprogesterone acetate, Fluorometholone, Melengestrol acetate, Dexamethasone, Prednisolone
CA1	176	0.5	4	Phenol, Sodium nitrite	Triclopyr, Triclopyr butotyl, p-Bromodiphenyl ether, 2-Fluoroacetamide, 1-Ethyl-2-methylbenzene
CA2	341	0.5	4	Celecoxib, Phenol	PharmaGSID_48509, Acenaphthylene, CP-105696, Aloe-emodin, 2-Fluoroacetamide
PTGS1	307	0.25	4	Indomethacin	SSR69071, 17alpha-Estradiol, Chlordane, Cetylpyridinium bromide, Zoxamide

Table Courtesy of Imran Shah

POTENTIAL APPLICATION FOR HTTR DATA IN CHEMICAL SAFETY DECISION MAKING PROCESSES

Biological Activity of APCRA Chemicals



Bioactivity Exposure Ratio (BER) Analysis Using HTTr



Bioactivity & Exposure Ratio Comparisons Using Reverse Dosimetry

- 5th Percentile of concentration-responsive gene (CRG) distribution
- Median BMD for most sensitive pathway

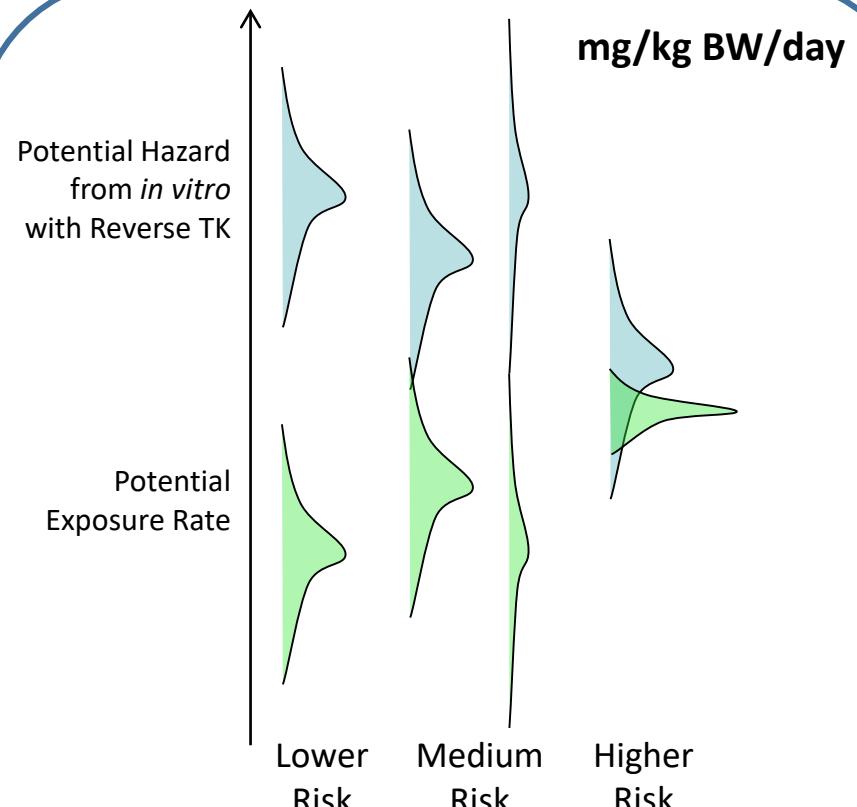
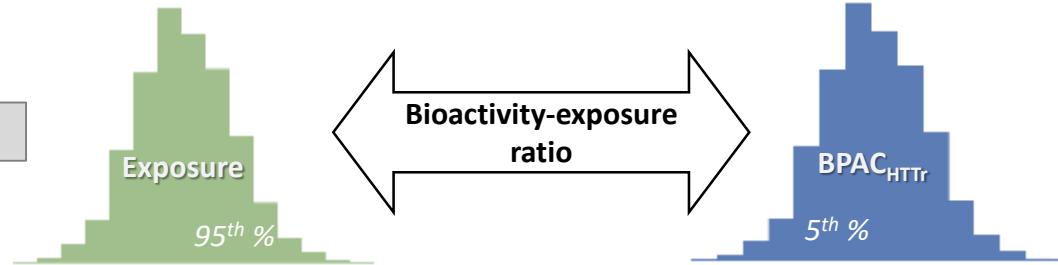
- Using httk v1.8 values for humans
- Default to a simple model with no partition coefficients and use of steady-state concentration.
- Assume 100% bioavailability and restrictive clearance.

HTTr BPAC
(μM)

Apply high-throughput toxicokinetics (httk) to get mg/kg/day

Potential Hazard from *in vitro* with Reverse TK

EPA - ExpoCast



High-throughput toxicokinetic (httk) modeling: Conversion of *in vitro* bioactivity to *in vivo* steady state concentration (C_{ss})

Reverse dosimetry: Conversion of predicted C_{ss} to an administered equivalent dose (AED)

Comparison of HTTr and ToxCast Potency Estimates

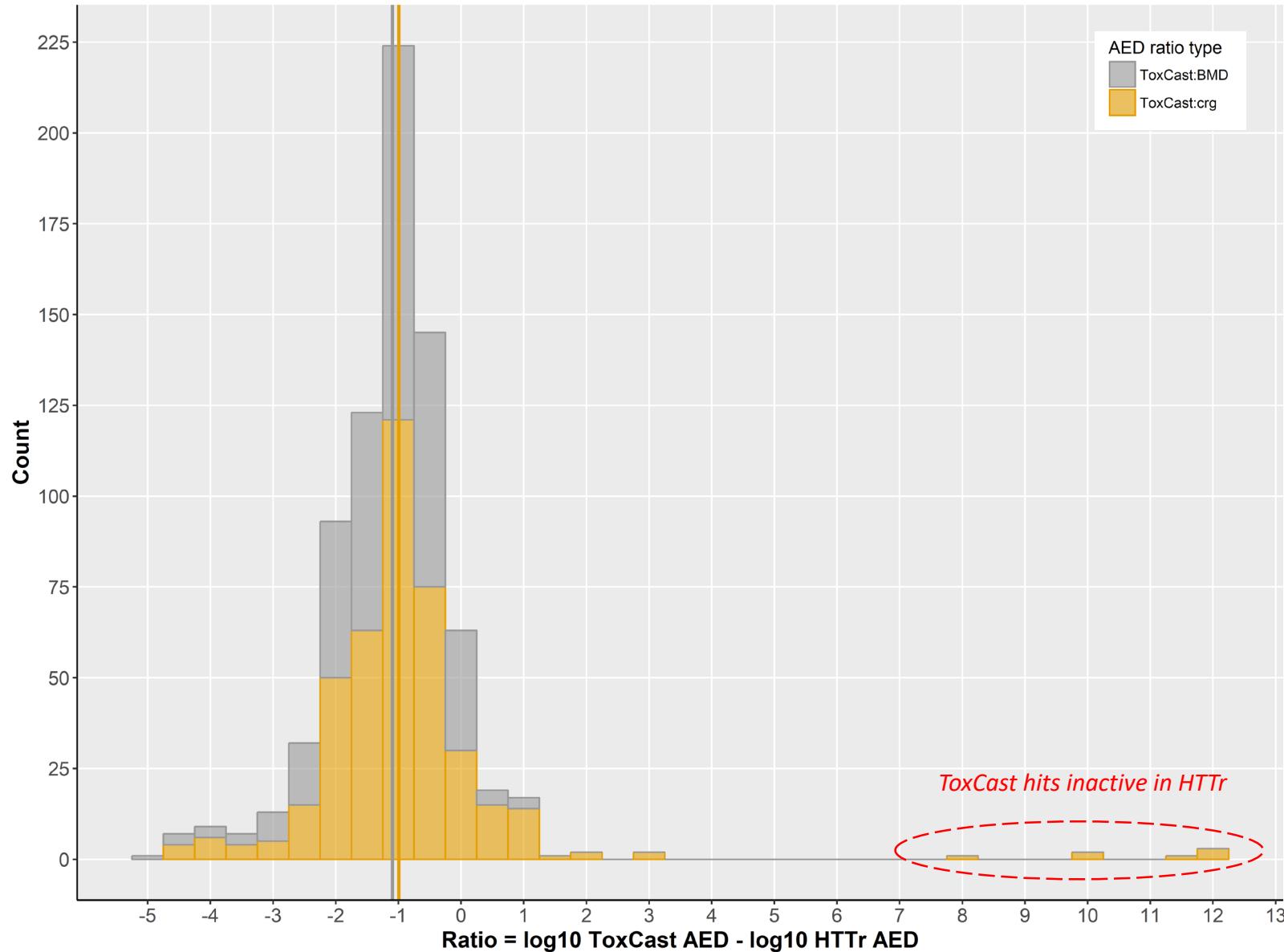


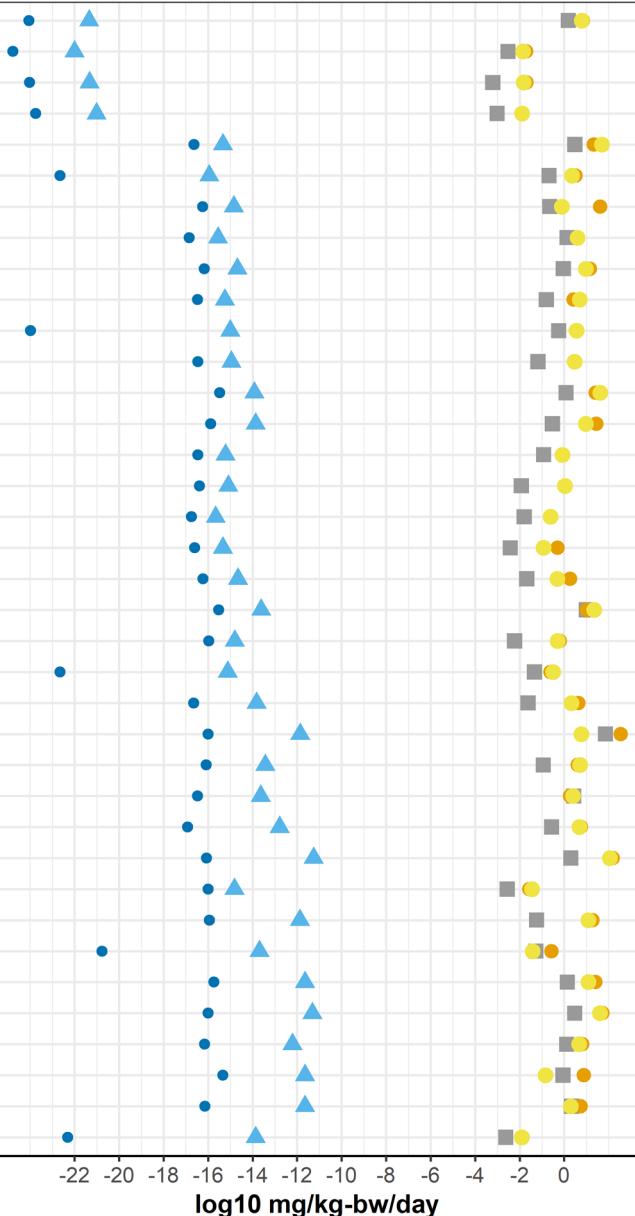
Figure Courtesy of Katie Paul-Friedman

BER Results

A

Dichlorodiphenyltrichloroethane
Aldrin
Endrin
Dieldrin
Acetochlor
Pentachlorophenol
Dimethoate
Carbosulfan
Azinphos-methyl
Methidathion
Heptachlor
Methyl parathion
Alachlor
Metolachlor
Chlorpyrifos-methyl
Parathion
Ethion
EPN
Fenitrothion
Pentachloronitrobenzene
Coumaphos
Endosulfan
Carbofuran
Trichlorfon
Fenthion
Dicrotophos
Propoxur
Sulfentrazone
Phorate
Isoxaben
Octrizole
Cyfluthrin
Imidacloprid
Nitrofurazone
Acephate
Thiophanate-methyl
p,p'-DDD

Lower Priority



B

Higher Priority

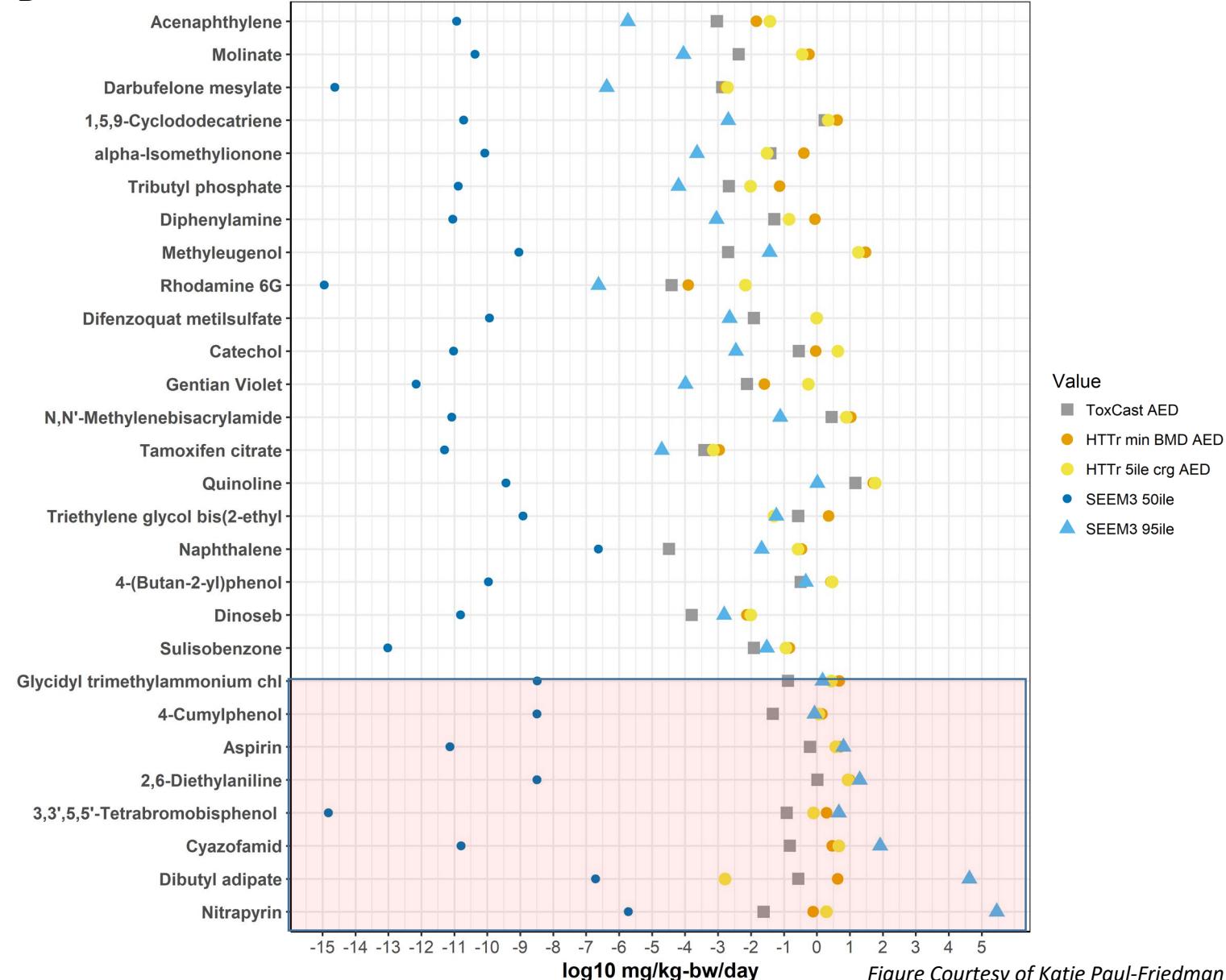


Figure Courtesy of Katie Paul-Friedman

HTTr Summary Slide

- **Technology:** Targeted RNA-Seq based HTTr is a promising platform for comprehensive and cost-effective evaluation of chemically-induced disruption of biological processes/pathways.
- **Workflow:** We have developed a standardized, scalable and portable workflow to generate large-scale HTTr data for thousands of chemicals.
- **Performance Standards:** The use of reference materials / QC standards on each plate enable development of performance standards for comparison within and across laboratories.
- **Concentration-Response Analysis:** Incorporation of concentration-response modeling into the analysis pipeline enables identification of transcriptional BPACs at the biological pathway/process level.
- **Bioactivity Exposure Ratio:** HTTr data may be used in combination with httk and ExpoCast estimates to identify chemicals with bioactivity thresholds in human relevant exposure ranges
- **MIE/MOA Identification:** Multiple analysis approaches are being investigated for identification of MIE/MOA. Target-centric signatures derived from annotated reference chemicals and machine learning techniques show promise for identification of putative MIE/MOAs.

Acknowledgments

NCCT:
Imran Shah
Katie Paul-Friedman
Clinton Willis
Woody Setzer
Derik Haggard
Richard Judson
Rusty Thomas

BioSpyder
Joel McComb
Bruce Seligmann
Jo Yeaaley
Milos Babic
Pete Shepard
Kyle LeBlanc
Jason Downing

Unilever:
Paul Carmichael
Andy White
Paul Russell
Sharon Scott
Sophie Malcomber

**National Toxicology
Program:**
Scott Auerbach

Sciome:
Jason Phillips



National Center for Computational Toxicology

